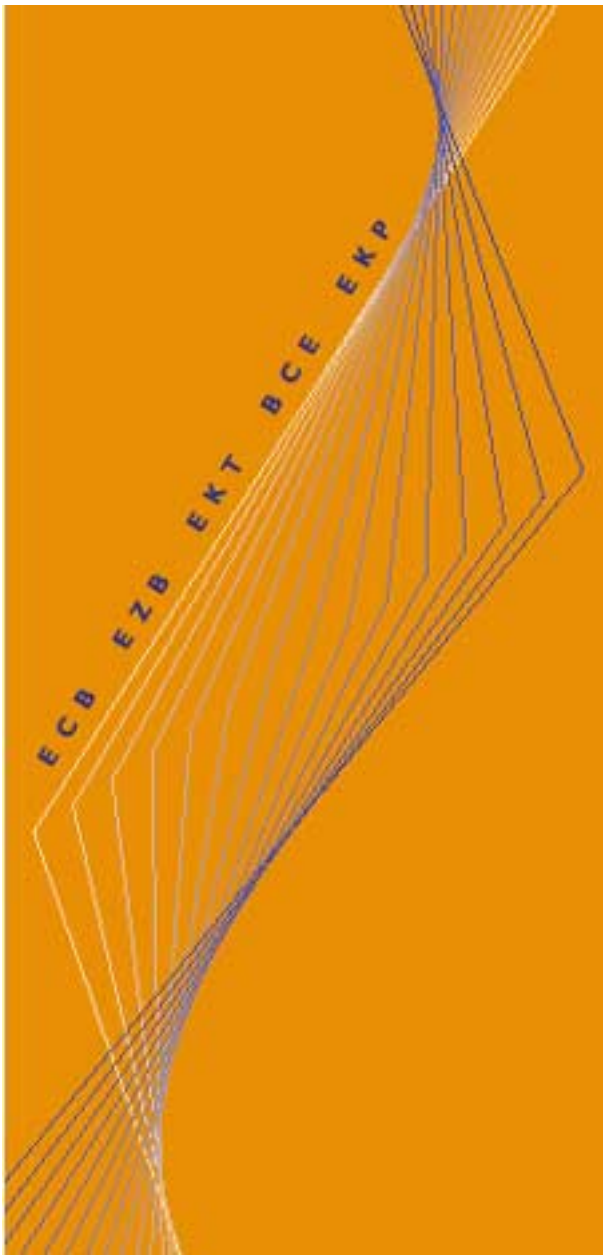
EUROPEAN CENTRAL BANK

WORKING PAPER SERIES



WORKING PAPER NO. 45

TESTING THE RANK OF THE HANKEL MATRIX: A STATISTICAL APPROACH

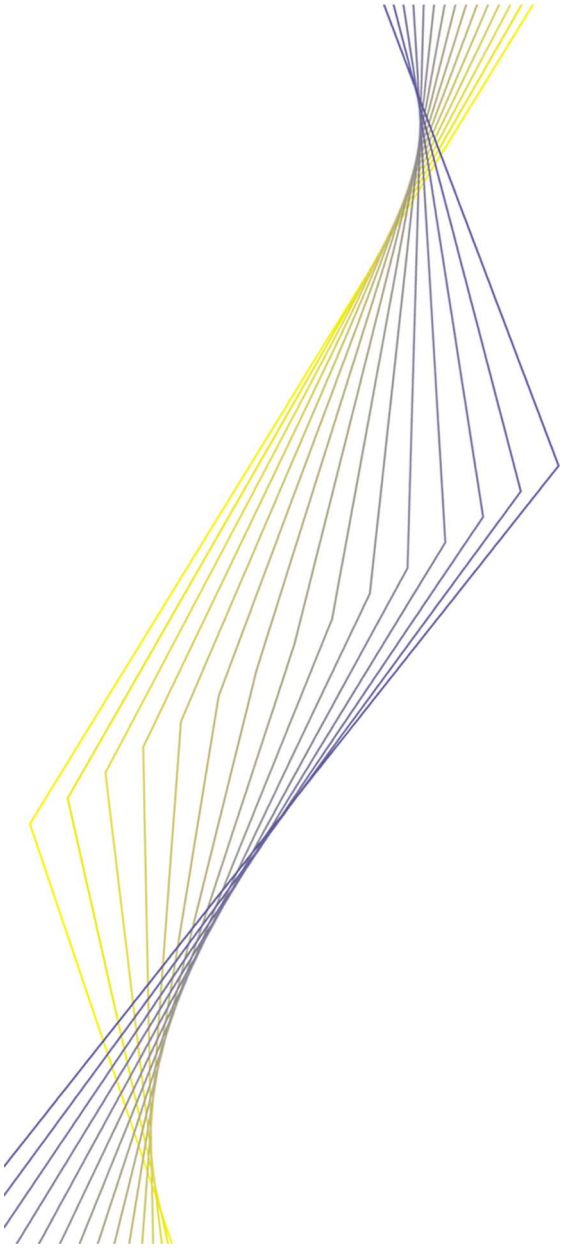
BY

GONZALO CAMBA-MENDEZ AND GEORGE KAPETANIOS

March 2001

EUROPEAN CENTRAL BANK

WORKING PAPER SERIES



WORKING PAPER NO. 45

TESTING THE RANK OF THE HANKEL MATRIX: A STATISTICAL APPROACH

BY

GONZALO CAMBA-MENDEZ* AND GEORGE KAPETANIOS**

March 2001

* *European Central Bank, Kaiserstrasse 29, D-60311, Frankfurt am Main, Germany, email: gonzalo.camba-mendez@ecb.int.*

** *National Institute of Economic and Social Research, 2 Dean Trench Street, Smith Square, SW1P 3HE, London, United Kingdom, email: gkapetanios@niesr.ac.uk.*

Address	**Kaiserstrasse 29** **D-60311 Frankfurt am Main** **Germany**
Postal address	**Postfach 16 03 19** **D-60066 Frankfurt am Main** **Germany**
Telephone	**+49 69 1344 0**
Internet	**http://www.ecb.int**
Fax	**+49 69 1344 6000**
Telex	**411 144 ecb d**

ISSN 1561-0810

Contents

Abstract

The rank of the Hankel matrix, corresponding to a system transfer function, is equal to the order of its minimal state space realization. The computation of the rank of the Hankel matrix is complicated by the fact that its block elements are rarely given exactly but are estimated instead. In this paper, we propose new statistical tests to determine the rank of the Hankel matrix. We also provide a Monte Carlo study on the reliability of these tests compared to existing procedures.

Keywords: Tests of Rank, Model Reduction, Hankel Operator, Monte Carlo.

1 Introduction

Multivariate time series procedures are very commonly used in Applied Economics studies. Atheoretical or unrestricted multivariate time series models are not meant to be a replacement to structural econometric models in the context of policy analysis or conditional forecasting, but they are usually the preferred choice in the context of unconditional forecasting. There are two main linear multivariate representations for a vector series: a VARMA representation and a State Space representation. Under certain parametric restrictions these representations are equivalent. In this paper we focus on the SS representation. State Space multivariate modeling has been used to model exchange rates, Dorfman (1997), economic interdependence between countries, Aoki (1987), and build a small macroeconometric model for the Dutch economy, Otter and Dal (1987).

A discrete stationary multivariate input-output system can be characterised by a system transfer function matrix $\boldsymbol{G}(z) = \sum_{i=1}^{\infty} \boldsymbol{G}_i z^{-i}$, where $\boldsymbol{G}_i$ are the impulse response matrices. Corresponding to the transfer function matrix $\boldsymbol{G}(z)$ above, the infinite dimensional Hankel matrix is defined as:

$$\begin{bmatrix} \boldsymbol{G}_1 & \boldsymbol{G}_2 & \boldsymbol{G}_3 & \cdots \\ \boldsymbol{G}_2 & \boldsymbol{G}_3 & \cdots & \cdots \\ \boldsymbol{G}_3 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{1}$$

Kronecker's theorem can be used to show that given a system transfer function matrix, $\boldsymbol{G}(z)$, the order of the system is equal to the rank of the Hankel matrix (see Kailath (1980)). The order of the system, $\boldsymbol{G}(z)$, is defined as the order of the minimal state-space realization, i.e. the minimal dimension of the state vector. The computation of the rank of the Hankel matrix is not an easy task, as it is unlikely that the impulse response matrices are given exactly, and in a majority of cases they are estimated. An alternative characterisation of the system in terms of a Hankel matrix of the covariances of the m-dimensional output vector, $\boldsymbol{y}_t$, is given by

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{\Delta}_1 & \boldsymbol{\Delta}_2 & \boldsymbol{\Delta}_3 & \cdots \\ \boldsymbol{\Delta}_2 & \boldsymbol{\Delta}_3 & \cdots & \cdots \\ \boldsymbol{\Delta}_3 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{2}$$

where $\boldsymbol{\Delta}_i$ denotes the i-th autocovariance of $\boldsymbol{y}_t$ (see, e.g., Aoki (1987, pp. 62-67)). Given $\boldsymbol{y}_t$, $t = 1, \ldots, T$, an estimator for a finite truncation of the Hankel matrix, $\boldsymbol{H}$, is denoted by $\hat{\boldsymbol{H}}$ and defined as follows[1]:

$$\hat{\boldsymbol{H}} = \frac{1}{T} \boldsymbol{\mathcal{Y}}^{+\prime} \boldsymbol{\mathcal{Y}}^{-} \tag{3}$$

[1]Throughout the note we use hats to indicate estimated quantities.

where $\boldsymbol{\mathcal{Y}}^+$ and $\boldsymbol{\mathcal{Y}}^-$ are defined as:

$$\boldsymbol{\mathcal{Y}}^- = \begin{bmatrix} \boldsymbol{y}_1' & 0 & \dots & 0 \\ \boldsymbol{y}_2' & \boldsymbol{y}_1' & \dots & 0 \\ \boldsymbol{y}_3' & \boldsymbol{y}_2' & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \boldsymbol{y}_{T-1}' & \boldsymbol{y}_{T-2}' & \dots & \boldsymbol{y}_{T-p}' \end{bmatrix} \quad \boldsymbol{\mathcal{Y}}^+ = \begin{bmatrix} \boldsymbol{y}_2' & \boldsymbol{y}_3' & \dots & \boldsymbol{y}_{k+1}' \\ \dots & \dots & \dots & \dots \\ \boldsymbol{y}_{T-2}' & \boldsymbol{y}_{T-1}' & \dots & 0 \\ \boldsymbol{y}_{T-1}' & \boldsymbol{y}_T' & \dots & 0 \\ \boldsymbol{y}_T' & 0 & \dots & 0 \end{bmatrix} \tag{4}$$

The above formulation enables the computation of the Hankel matrix without the need to calculate covariance matrices. The rank of $\boldsymbol{H}$ gives the number of states required to approximate the autocovariance sequence arranged in $\boldsymbol{H}$. For more details see Kung and Lin (1981), Kailath (1980) and Aoki (1987). In this paper, we propose new statistical tests to determine the rank of the Hankel matrix. We also provide a Monte Carlo study on the reliability of these tests compared to existing procedures.

Section 2 presents different procedures to determine the rank of the Hankel matrix. Two of those procedures are formal statistical tests, one is a rule of thumb with unknown statistical properties, and a further two are based on information theoretic arguments. Section 3 describes the design of the Monte Carlo simulations computed to assess the performance of the procedures. Section 4 presents the results of the Monte Carlo experiments, and section 5 concludes.

2 Testing the rank of the Hankel matrix $\boldsymbol{H}$

The rank of $\boldsymbol{H}$ is the focus of the investigation. Under the assumption of stationarity of the input-output multivariate system, it can be shown[2] that $\sqrt{T}vec(\hat{\boldsymbol{H}} - \boldsymbol{H})$ is asymptotically distributed as $N(\boldsymbol{0}, \boldsymbol{V})$. One of the statistical tests of rank requires an estimator of the covariance matrix of $\sqrt{T}vec(\hat{\boldsymbol{H}} - \boldsymbol{H})$, $\boldsymbol{V}$. A consistent[3] estimator of $\boldsymbol{V}$ is given by: $\hat{\boldsymbol{V}} = T^{-1}\sum_{t=k}^{T-k}\left(vec(\boldsymbol{\mathcal{Z}}_t) - vec(\hat{\boldsymbol{H}})\right)\left(vec(\boldsymbol{\mathcal{Z}}_t) - vec(\hat{\boldsymbol{H}})\right)'$ where $\boldsymbol{\mathcal{Z}}_t = \left((\boldsymbol{y}_t\boldsymbol{\mathcal{Y}}_{t-1}^-)', (\boldsymbol{y}_t\boldsymbol{\mathcal{Y}}_{t-2}^-)', \dots (\boldsymbol{y}_t\boldsymbol{\mathcal{Y}}_{t-k}^-)'\right)'$ and $\boldsymbol{\mathcal{Y}}_t^-$ denotes the t-th row of $\boldsymbol{\mathcal{Y}}^-$. Note that by construction both the asymptotic covariance matrix and its finite sample estimator are of reduced rank.

The procedures to test for the rank of $\boldsymbol{H}$, denoted as $rk(\boldsymbol{H})$, consider the following hypothesis[4], $H_0 : rk(\boldsymbol{H}) = r^*$ against the alternative hypothesis, $H_1 : rk(\boldsymbol{H}) > r^*$. Starting with the null hypothesis of $r^* = 1$, a sequence of tests is performed. If the null is rejected, r^* is augmented by one and the test is repeated. When the null cannot be rejected, r^* is adopted as the estimate of the rank of $\boldsymbol{H}$. For each test, a test statistic is constructed. This

[2] See e.g. Brockwell and Davis (1991, Ch. 7)

[3] An estimator for a quantity is consistent if the estimate tends to the true value of the quantity in probability, asymptotically.

[4] Referred to as the null hypothesis in the statistical literature.

statistic is shown to follow a given statistical distribution under the null hypothesis. When a comparison of the statistic with chosen percentiles of the distribution indicate that the statistic is unlikely to be a realisation of that distribution the null hypothesis is rejected. Usually in statistical applications using one-sided statistical tests, the chosen percentile of the statistical distribution, denoted by $1-\alpha$, is the 95% percentile, where α is referred to as the significance level and it is assumed that large values of the test statistic provide evidence for rejecting the null hypothesis. Nevertheless, the rank estimate provided by this approach will not converge in probability to the true value of the rank of the Hankel matrix, denoted by r^0. The reason is that even if the null hypothesis tested is true, the testing procedure will reject it with probability α. The rank estimate will converge to its true value, r^0, as T goes to infinity, if α is made to depend on T and goes to zero as T goes to infinity but not faster than a given rate. We denote this α by α_T, where the subscript T now denotes dependence of the significance level on the sample size. Hosoya (1989) shows that if α_T goes to zero as the sample size T goes to infinity and also $\lim_{T\to\infty} \ln \alpha_T/T = 0$, then the rank estimate provided by the sequential testing procedure will converge in probability to r^0, see also Cragg and Donald (1997).

2.1 Cragg and Donald (1996)

The procedure proposed by Cragg and Donald (1996) is based on the transformation of the matrix $\boldsymbol{H}$ using Gaussian elimination with complete pivoting[5]. r^* steps of Gaussian elimination with full pivoting on matrix $\boldsymbol{H}$ amounts to the following operations:

$$\boldsymbol{Q}_{r^*}\boldsymbol{R}_{r^*}\boldsymbol{Q}_{r^*-1}\boldsymbol{R}_{r^*-1}\ldots\boldsymbol{Q}_1\boldsymbol{R}_1\boldsymbol{H}\boldsymbol{C}_1\ldots\boldsymbol{C}_{r^*-1}\boldsymbol{C}_{r^*} = \begin{bmatrix} \boldsymbol{H}_{11}(r^*) & \boldsymbol{H}_{12}(r^*) \\ \boldsymbol{0} & \boldsymbol{H}_{22}(r^*) \end{bmatrix}$$

where $\boldsymbol{R}_i$ and $\boldsymbol{C}_i$ are pivoting matrices for step i and $\boldsymbol{Q}_i$ are Gauss transformation matrices. The pivoting matrices used to perform the first r^* steps of Gaussian elimininations are applied to $\boldsymbol{H}$ to obtain the following relation

$$\boldsymbol{R}_{r^*}\boldsymbol{R}_{r^*-1}\ldots\boldsymbol{R}_1\boldsymbol{H}\boldsymbol{C}_1\ldots\boldsymbol{C}_{r^*-1}\boldsymbol{C}_{r^*} = \boldsymbol{R}\boldsymbol{H}\boldsymbol{C} = \boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_{11}(r^*) & \boldsymbol{F}_{12}(r^*) \\ \boldsymbol{F}_{21}(r^*) & \boldsymbol{F}_{22}(r^*) \end{bmatrix} \tag{5}$$

where $\boldsymbol{F}$ is partitioned accordingly, i.e. $\boldsymbol{F}_{11}(r^*)$ is of dimension $r^* \times r^*$. Note that in this case $\boldsymbol{F}_{11}(r^*)$ has full rank, under the null hypothesis that $r^* = r^0$. It then follows, (see Cragg and Donald (1996)), that $\boldsymbol{F}_{22}(r^*) - \boldsymbol{F}_{21}(r^*)\boldsymbol{F}_{11}^{-1}(r^*)\boldsymbol{F}_{12}(r^*) = 0$. The estimated counterpart of the above relation, i.e. $\hat{\boldsymbol{F}}_{22} - \hat{\boldsymbol{F}}_{21}\hat{\boldsymbol{F}}_{11}^{-1}\hat{\boldsymbol{F}}_{12} = \hat{\boldsymbol{\Lambda}}_{22}(r^*)$, may be used as a test statistic of the hypothesis that the rank of $\boldsymbol{H}$ is r^*. Under regularity conditions, including the requirement

[5]For details on Gaussian elimination with complete pivoting see Cragg and Donald (1996) or Golub and Loan (1983).

that the covariance matrix of the asymptotically normally distributed matrix $\sqrt{T}vec(\hat{\boldsymbol{H}}-\boldsymbol{H})$ has full rank, the following result can be shown, under H_0.

$$\sqrt{T}vec(\hat{\boldsymbol{\Lambda}}_{22}(r^*)) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{\Gamma V \Gamma}')$$

where $\boldsymbol{\Gamma} = \boldsymbol{\Phi}_2 \otimes \boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_1 = \left[-\hat{\boldsymbol{F}}_{21}\hat{\boldsymbol{F}}_{11}^{-1} \;\; \boldsymbol{I}_{mk-r^*}\right]\boldsymbol{R}$, $\boldsymbol{\Phi}_2 = \left[-\hat{\boldsymbol{F}}_{12}'\hat{\boldsymbol{F}}_{11}^{-1'} \;\; \boldsymbol{I}_{mp-r^*}\right]\boldsymbol{C}'$ and $\xrightarrow{d}$ denotes convergence in distribution. Then,

$$\hat{\xi} = T vec\ \hat{\boldsymbol{\Lambda}}_{22}(r^*)'(\hat{\boldsymbol{\Gamma}}\hat{\boldsymbol{V}}\hat{\boldsymbol{\Gamma}}')^{-1} vec\ \hat{\boldsymbol{\Lambda}}_{22}(r^*) \xrightarrow{d} \chi^2_{(mk-r^*)(mp-r^*)}$$

where $\hat{\boldsymbol{\Gamma}}$ and $\hat{\boldsymbol{V}}$ are the sample estimates of $\boldsymbol{\Gamma}$ and $\boldsymbol{V}$ and χ^2_l denotes the χ^2 distribution with l degrees of freedom.

The procedure uses the inverse of the covariance matrix of the Hankel matrix. However, this is not available. Therefore, we modify the existing procedure and use a generalised inverse instead. We distinguish between two cases. In the first case, the rank of the estimated covariance matrix is, by construction or otherwise, equal to that of the asymptotic covariance matrix. This is the case for the estimated covariance matrix of the Hankel matrix described above. In the second case, the rank of the estimated covariance matrix is larger than that of the asymptotic covariance matrix. Nevertheless, the estimated covariance is a consistent estimate. The following proposition covers both cases.

Proposition 1 *Let $\sqrt{T}vec(\hat{\boldsymbol{H}}-\boldsymbol{H})$ be asymptotically distributed as $N(\boldsymbol{0}, \boldsymbol{V})$ where $\boldsymbol{V}$ can be of reduced rank and $\hat{\boldsymbol{H}}$ is of full rank. Let $\hat{\boldsymbol{V}}$ be a consistent estimate of $\boldsymbol{V}$. Further, assume that $\boldsymbol{R}_i$ and $\boldsymbol{C}_i$, $i = 1, \ldots, r^*$ in (5) are uniquely defined and, without loss of generality, that $\boldsymbol{R}_i = \boldsymbol{I}$, $\boldsymbol{C}_i = \boldsymbol{I}$, $i = 1, \ldots, r^*$. Then, the following result holds under the null $H_0 : r = r^*$,*

$$\hat{\xi} = T vec\ \hat{\boldsymbol{\Lambda}}_{22}(r^*)'(\hat{\boldsymbol{\Gamma}}\hat{\boldsymbol{V}}\hat{\boldsymbol{\Gamma}}')^{+} vec\ \hat{\boldsymbol{\Lambda}}_{22}(r^*) \xrightarrow{d} \chi^2_{\beta}$$

if additionally the rank of $\boldsymbol{V}$ is known and

$$rk[\hat{\boldsymbol{V}}] = rk[\boldsymbol{V}], \quad \forall T \tag{6}$$

where $^+$ denotes the Moore-Penrose inverse of a matrix, and the number of degrees of freedom β is given by the minimum between the number of rows in $\hat{\boldsymbol{\Gamma}}$ and the rank of $\hat{\boldsymbol{V}}$; i.e. $\min\{(mp-r^*)\times(mk-r^*), (k+p-1)*m\}$.

Proof: Given the remarks following Assumption 2 of Cragg and Donald (1996) it suffices to consider the case where $\boldsymbol{\Gamma}$ is a function of submatrices of $\boldsymbol{H}$ as given in page 1304 of Cragg and Donald (1996). Following Andrews (1987), the Proposition holds if $(\hat{\boldsymbol{\Gamma}}\hat{\boldsymbol{V}}\hat{\boldsymbol{\Gamma}}')^{+} \xrightarrow{p} (\boldsymbol{\Gamma V \Gamma}')^{+}$. But by (6), $\hat{\boldsymbol{V}} = \boldsymbol{\Upsilon}\hat{\boldsymbol{\Xi}}\boldsymbol{\Upsilon}'$ where $\hat{\boldsymbol{\Xi}}$ is the estimate of the variance of the distinct elements of $\sqrt{T}vec(\hat{\boldsymbol{H}} - \boldsymbol{H})$ and is of full rank and $\boldsymbol{\Upsilon}$ is a matrix of ones and zeros which duplicates

elements of $\boldsymbol{\Xi}$. Therefore, $\hat{\boldsymbol{\Gamma}}\hat{\boldsymbol{V}}\hat{\boldsymbol{\Gamma}}' = \hat{\boldsymbol{\Gamma}}\boldsymbol{\Upsilon}\hat{\boldsymbol{\Xi}}\boldsymbol{\Upsilon}'\hat{\boldsymbol{\Gamma}}' = \hat{\boldsymbol{\Pi}}\hat{\boldsymbol{V}}\hat{\boldsymbol{\Pi}}'$, where $\hat{\boldsymbol{\Pi}} = \hat{\boldsymbol{\Gamma}}\boldsymbol{\Upsilon}$. Note that $\hat{\boldsymbol{\Pi}}$ is simply obtained from $\hat{\boldsymbol{\Gamma}}$ by adding the columns of $\hat{\boldsymbol{\Gamma}}$ corresponding to the identical elements of $\hat{\boldsymbol{H}}$. For all $T > M$, where M is large number, $\hat{\boldsymbol{\Pi}}$ is of full row or column rank (whichever is smaller). This would not be the case only if two columns of $\hat{\boldsymbol{\Gamma}}$ added up to a vector of zeros. If that were the case, linear combinations of distinct elements of $\hat{\boldsymbol{H}}$ would be converging to their true values at a rate faster than $T^{1/2}$ which is not allowed by the central limit theorem for stationary processes. Therefore, the Propostion holds by Theorem 2 of Andrews (1987).

Note that, as shown in Cragg and Donald (1996), the assumption concerning the identification of $\boldsymbol{R}_i$ and $\boldsymbol{C}_i$, $i = 1, \ldots, r^*$, is not neccesary for the above result to hold. Trivially, condition (6) is satisfied for $\hat{\boldsymbol{V}}$ since its rank is equal to that of $\boldsymbol{V}$ for all T. We will denote the test by GE_g.

2.2 Bartlett (1947)

An alternative method to estimate the rank of the Hankel matrix is based on the computation of canonical correlations. A well known result in canonical correlation analysis is that given two random vector series $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ of dimensions k and p respectively, the rank of the covariance matrix between those two random vectors is equal to the number of nonzero canonical correlations, see Anderson (1984) for further details. The Hankel matrix, defined in (3), is the covariance matrix between two random vectors $\boldsymbol{y}_t^+$ and $\boldsymbol{y}_t^-$ defined as $\boldsymbol{y}_t^+ = \left(\boldsymbol{y}_{t+1}', \ldots, \boldsymbol{y}_{t+k}'\right)'$ and $\boldsymbol{y}_t^- = \left(\boldsymbol{y}_t', \ldots, \boldsymbol{y}_{t-p+1}'\right)'$. Compute the QR decomposition of the matrices $\boldsymbol{\mathcal{Y}}^+$ and $\boldsymbol{\mathcal{Y}}^-$ given in (4), i.e. $\boldsymbol{\mathcal{Y}}^+ = \boldsymbol{Q}^+\boldsymbol{R}^+$ and $\boldsymbol{\mathcal{Y}}^- = \boldsymbol{Q}^-\boldsymbol{R}^-$. The canonical correlations between the vectors $\boldsymbol{y}_t^+$ and $\boldsymbol{y}_t^-$, are the singular values of $\boldsymbol{Q}^{+'}\boldsymbol{Q}^-$. We denote the canonical correlations as ρ_i, $i = 1, \ldots, \min(k, p)m$. Bartlett (1947) provided a likelihood ratio criterion for testing the null hypothesis that the last $r_{\min(\mathrm{k,p})m} - r^* - 1$ canonical correlations are zero, i.e., $H_{r^*} : \rho_{r^*+1} = \cdots = \rho_{\min(\mathrm{k,p})m} = 0$. Under the null hypothesis and assuming stationarity of the input-output multivariate system

$$BA = \left[\frac{m(k+p)+1}{2} - T\right] \ln \prod_{i=r^*+1}^{\min(\mathrm{k,p})m} \left(1 - \hat{\rho}_i^2\right) \xrightarrow{d} \chi^2_{(mk-r^*)\times(mp-r^*)}$$

Bartlett's test was developed under independence and normality assumptions, but his result remains valid asymptotically following arguments by Kohn (1979) on the likelihood ratio tests for dependent observations. We note that the Cragg and Donald procedure is a more general procedure for determining the rank of a matrix since it only requires that an estimate of that matrix exists having a normal asymptotic distribution with a covariance matrix whose rank is known. The Barlett procedure is applicable in this context because the problem can be recast in terms of canonical correlations.

2.3 Akaike (1976)

Akaike (1974) and Akaike (1976) showed that the number of linearly independent components of the projections of the previously defined $\boldsymbol{y}_t^+$ onto the linear space spanned by the components of $\boldsymbol{y}_t^-$ is identical to the number of nonzero canonical correlations between $\boldsymbol{y}_t^+$ and $\boldsymbol{y}_t^-$. When $\boldsymbol{y}_t$ is Gaussian, canonical correlation analysis between $\boldsymbol{y}_t^+$ and $\boldsymbol{y}_t^-$ is equivalent to maximum likelihood estimation of the linear model: $\boldsymbol{y}_t^+ = \boldsymbol{\Psi}\boldsymbol{y}_t^- + \boldsymbol{\varepsilon}_t$, see Anderson (1984). The number of free parameters for this model is: $F(r^*) = \{[s^+(s^+ + 1)]/2\} + \{[s^-(s^- + 1)]/2\} + r^*(s^+ + s^- - r^*)$ where s^+ denotes the dimension of the vector $\boldsymbol{y}_t^+$ and s^- denotes the dimension of $\boldsymbol{y}_t^-$. The first two terms are the number of free parameters of the covariance matrices of $\boldsymbol{y}_t^+$ and $\boldsymbol{y}_t^-$ respectively, and the last term gives the number of free parameters in matrix $\boldsymbol{\Psi}$. Akaike (1976) defined an information criterion for model fitting, and by extension rank determination, as:

$$AIC(r^*) = T \ln \prod_{i=1}^{r^*} (1 - \hat{\rho}_i^2) + 2F(r^*)$$

where $\hat{\rho}_i$ are the estimated canonical correlation coefficients previously defined. This criterion penalises models with a large number of parameters, and by extension large rank, and favours parsimonious representations. Note that, as discussed in Anderson (1984, pp. 499), when $\rho_i = 0$ then $\hat{\rho}_i^2 = O_p(T^{-1})$, implying that $\ln(1 - \hat{\rho}_i^2) = O_p(T^{-1})$ where $O_p(.)$ denotes order in probability. This suggests that there is a positive probability that AIC will be minimised for some $r^* > r^0$ since the probability that $T \sum_{i=r^0+1}^{r^*} \ln(1 - \hat{\rho}_i^2) < 2(F(r^0) - F(r^*))$ is greater than zero. Therefore, the estimated rank will not converge in probability to r^0 when AIC is used.

2.4 Schwarz (1978)

Schwarz (1978) suggested an alternative penalty on increasing the number of parameters, and in the present paper we explore also the performance of this criterion in searching for the rank. The information criterion is:

$$BIC(r^*) = T \ln \prod_{i=1}^{r^*} (1 - \hat{\rho}_i^2) + \ln(T) F(r^*)$$

The penalty used by BIC is much more severe than that used by AIC. In fact, it is easy to see that the rank estimate obtained by BIC will converge in probability to r^0. Nevertheless, BIC is likely to underestimate the rank in small samples.

2.5 Rule of Thumb

This method, suggested in Aoki (1987), is based on the singular value decomposition of the estimated Hankel matrix, $\hat{\boldsymbol{H}}$, given by $\hat{\boldsymbol{H}} = \hat{\boldsymbol{U}}\hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{V}}$; where $\hat{\boldsymbol{U}}$ and $\hat{\boldsymbol{V}}$ are orthogonal matrices, and $\hat{\boldsymbol{\Sigma}}$ is a rectangular matrix of zero nondiagonal elements, with diagonal elements equal to the singular values of $\hat{\boldsymbol{H}}$, $\hat{\sigma}_1 \geq \hat{\sigma}_2 \geq \ldots \geq \hat{\sigma}_{\min(\mathrm{k,p})m}$. The rule of thumb suggests that the rank is equal to r^* if

$$\frac{\hat{\sigma}_{r^*+1}}{\hat{\sigma}_1} \leq \frac{1}{\sqrt{T}} \quad \text{and} \quad \frac{\hat{\sigma}_i}{\hat{\sigma}_1} > \frac{1}{\sqrt{T}}, \quad i = 1, \ldots, r^* \tag{7}$$

where T is the number of observations available for the vector series $\boldsymbol{y}_t$. Clearly, the number of nonzero singular values of $\boldsymbol{H}$ will be equal to the rank of $\boldsymbol{H}$. Since the Hankel matrix is the covariance matrix between $\boldsymbol{y}_t^+$ and $\boldsymbol{y}_t^-$, its singular values will also be the canonical correlations[6] between $\boldsymbol{y}_t^+$ and $\boldsymbol{y}_t^-$. Therefore, as discussed in subsection 2.3 above, the estimated singular values will be $O_p(T^{-1/2})$. This remark is the motivation behind (7) as a rule of thumb.

3 Monte Carlo Design

We concentrate on the state space model

$$\boldsymbol{y}_t = \boldsymbol{C}\boldsymbol{s}_t + \boldsymbol{e}_t, \quad \boldsymbol{s}_{t+1} = \boldsymbol{A}\boldsymbol{s}_t + \boldsymbol{B}\boldsymbol{e}_t \tag{8}$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are $r \times r$, $r \times m$ and $m \times r$ parameter matrices respectively, $\boldsymbol{s}_t$ is an r-vector of unobservable state variables, and $\boldsymbol{e}_t$ is an m-vector of random variables with mean zero and positive definite covariance matrix $\boldsymbol{\Omega}$.

The dimension of the vector series $\boldsymbol{y}_t$ is fixed to three. The rank of the Hankel matrix is equal to the dimension of the state vector $\boldsymbol{s}_t$ which is fixed to three as well. Matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{\Omega}$ are built as follows. $\boldsymbol{B}$ and $\boldsymbol{C}$ are (3×3) matrices of values drawn from a standard normal distribution; $\boldsymbol{\Omega}$ is fixed to an identity matrix of dimension (3×3). Note that $\boldsymbol{A}$ is a key matrix to explain the dynamics of the series; the degree of persistence of shocks will depend on the eigenvalues of $\boldsymbol{A}$. To control the experiment for this, we have chosen to build $\boldsymbol{A} = \tilde{\boldsymbol{E}}\boldsymbol{\Lambda}\tilde{\boldsymbol{E}}'$. $\boldsymbol{\Lambda}$ is a 3×3 quasi upper triangular matrix; the last element of the diagonal corresponds to the modulus assigned to that experiment, and the 2×2 block matrix in the left upper corner is computed in such a way that the modulus of the complex pair of eigenvalues of this 2×2 block is also equal to the modulus assigned to the eigenvalues of that experiment; the remaining values are fixed to a value of one. $\tilde{\boldsymbol{E}}$ is an orthonormal matrix generated from a standard normal matrix using Gram-Schmidt orthogonalisation.

[6] Assuming for simplicity that both $\boldsymbol{y}_t^+$ and $\boldsymbol{y}_t^-$ have identity covariance matrices.

For the Monte Carlo experiments presented below, 3 different moduli making 3 alternative experiments will be used, namely 0.8, 0.4 and 0.2.

Using $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and random normal disturbances generated by a random number generator with an identity covariance matrix[7] a sample from a process following (8) is obtained. The sample sizes considered are: 250, 500 and 2000. For each simulated sample, 200 initial observations have been discarded, to minimise the effect of starting values which are set to zero. Each simulated sample is normalised to have zero mean and an identity covariance matrix. For each Monte Carlo experiment 2000 replications have been carried out. For all experiments, the Hankel matrix is computed for $k = p$. We consider $k = 2, 3, 4$. Finally, it is of interest to study the performance of the tests when the output series is observed with noise. Two sets of data are used, one is generated from (8) above, and the second is obtained by adding an extra noise component from a standard normal distribution to the output generated from (8). We refer to these experiments as *Noise Free Data* and *Noisy Data* respectively. Following Cragg and Donald (1997), we have specified $\alpha_T = \kappa / \ln(T)$ where κ is chosen so that $\alpha_{50} = 0.05$. All computations were carried out using the GAUSS programming language. We note that the computational demands of the statistical testing procedures are small. Indicatively, on a 400Mhz IBM-compatible PC running Windows 98, for $T = 250$, a system of dimension 3 and $k = p = 4$, the GE_g procedure took 1.43 sec to provide a rank estimate whereas the BA procedure took 0.93 sec.

4 Monte Carlo Results on Rank Estimation

Tables 1 and 2 present the mean and root mean square error (RMSE) of the estimated ranks over the Monte Carlo simulations for all experiments. Additionally, tables 3 and 4 display the distribution of the estimated rank. The results on RMSE are in accordance with expectations. Performance worsens when the dimension of the Hankel matrix increases, when data are contaminated with added noise, when experiments with smaller eigenvalue moduli are considered, and when the sample size is small. But there are some interesting disparities in the relative performance of the alternative procedures.

Both statistical tests and procedures based on information criteria perform much better than the rule of thumb. When looking at the size of the RMSE, it is clear that the performance of the rule of thumb is very bad in a number of cases. In terms of RMSE, the rule of thumb never performs better than all the other procedures, in the 54 experiments run. Unlike the statistical testing procedures, its performance worsens as the number of observations increases for a number of cases indicating that asymptotically this procedure is

[7] Limited experimentation with alternative covariance matrices has been undertaken. The main conclusions were not affected.

flawed.

The performance of the GE_g test is very good when $k = 2, 3$. This result is relatively unaffected by the number of observations, whether the data are contaminated with added noise or not, or how large the moduli of the eigenvalues are. The bad performance of the GE_g for Hankel matrices with $k = 4$ should come as no surprise. This test relies on the estimated covariance matrix, $\hat{\boldsymbol{V}}$, which is a 144×144 matrix for $k = 4$. The large dimension of the covariance matrix implies that it is likely to be poorly estimated.

BA and AIC appear to be less sensitive than the other procedures to all dimensions explored in the Monte Carlo exercises. In particular, both BA and AIC are much more robust to large numbers of blocks in the Hankel matrix and to small samples. Of these two procedures, the performance of BA, both in terms of mean and RMSE, is usually much better. The performance of B is either the best or is close to the best for every experiment.

The performance of BIC doesn't deteriorate much, when the number of blocks in the Hankel matrix increases. Nevertheless, it appears to be more sensitive than the other procedures to sample size. For example, for a sample size of 250 BIC is the best in terms of RMSE only in two out of eighteen experiments, whereas it is the best in eight out of eighteen experiments when the sample size is 2000.

5 Conclusion

Alternative methods to test the rank of a Hankel matrix have been described. A recently proposed statistical test of rank has been extended to deal with the peculiarities of this problem. The performance of the alternative procedures has been studied. This study was conducted by means of a Monte Carlo exercise which served to assess the sensitivity of different procedures to four dimensions: i) Number of blocks in the Hankel matrix, ii) sample size, iii) size of the moduli of the eigenvalues of matrix $\boldsymbol{A}$ in model (8) and iv) the presence of added noise in the output series. BA and AIC are less sensitive than the other procedures to these dimensions. Statistical tests of rank like the GE_g and B and information criteria like AIC and BIC were shown to have a better performance than the rule of thumb whose statistical properties are unknown. GE_g performs well when the number of blocks in the Hankel matrix is small, but its performance deteriorates for a large number of blocks. The performance of BIC is sensitive to the sample size.

References

AKAIKE, H. (1974): "Stochastic Theory of Minimal Realisations," *IEEE Transactions on Automatic Control*, AC-19(6), 667–674.

——— (1976): "Canonical Correlation Analysis of Time Series and the Use of an Information Criterion," in *System Identification*, ed. by R. Mehra, and D. Lainiotis. Academic Press.

ANDERSON, T. W. (1984): *Introduction to Multivariate Statistical Analysis*. John Wiley and Sons, New York.

ANDREWS, D. W. K. (1987): "Asymptotic Results for Generalised Wald Tests," *Econometric Theory*, 3, 348–358.

AOKI, M. (1987): *State Space Modelling of Time Series*. Springer-Verlag.

BARTLETT, M. (1947): "Multivariate Analysis," *Journal of the Royal Statistical Society, Series B*, 9, 176–197.

BROCKWELL, P. J., AND R. A. DAVIS (1991): *Time Series: Theory and Methods*. Springer Series in Statistics, Springer, New York.

CRAGG, J. G., AND S. G. DONALD (1996): "On the Asymptotic Properties of LDU-Based Tests of the Rank of a Matrix," *Journal of the American Statistical Association*, 91(435), 1301–1309.

——— (1997): "Inferring the Rank of a Matrix," *Journal of Econometrics*, 76(1–2), 223–250.

DORFMAN, J. H. (1997): "Competing exchange rate models: a state space model vs structural time series alternatives," in *Applications of computer aided time series modeling. Lecture Notes in Statistics, Vol. 119*, ed. by M. Aoki, and A. M. Havenner. Springer, New York.

GOLUB, G. H., AND C. F. V. LOAN (1983): *Matrix Computation*. North Oxford Academic.

HOSOYA, Y. (1989): "Hierarchical Statistical Models and a Generalised Likelihood Ratio Test," *Journal of the Royal Statistical Society B*, 51(3), 435–448.

KAILATH, T. (1980): *Linear Systems*. Englewood Cliffs, N. J.: Prentice-Hall.

KOHN, R. (1979): "Asymptotic Estimation and Hypothesis Testing Results for Vector Linear Time Series Models," *Econometrica*, 47(4), 1005–1030.

KUNG, S., AND D. W. LIN (1981): "Optimal Hankel-Norm model Reductions: Multivariable Systems," *IEEE Transactions on Automatic Control*, AC-26(4), 832–852.

OTTER, P. W., AND R. V. DAL (1987): "State Space and Distributed Lag Modelling of Dynamic Economic Processes Based on Singular Value Decompositions: with Applications to the Dutch Economy," *Annales D'Economie et de Statistique*, 6/7, 253–277.

SCHWARZ, G. (1978): "Estimating the Dimension of a Model," *Annals of Statistics*, 6, 461–464.

Table 1

Mean and RMSE of Estimated Rank. Noise Free Data.

mod [a]	k	Tests	250		500		2000	
			$\overline{x}$	$\Re$	$\overline{x}$	$\Re$	$\overline{x}$	$\Re$
0.8	2	GE_g	2.96	0.25	2.99	0.17	3.00	0.09
		BA	3.04	0.31	3.06	0.29	3.06	0.26
		AIC	3.18	0.51	3.19	0.50	3.20	0.50
		BIC	2.92	0.29	2.97	0.18	2.99	0.10
		$Thumb$	2.81	0.49	2.90	0.36	2.98	0.22
0.4	2	GE_g	2.87	0.50	2.97	0.41	3.01	0.24
		BA	2.86	0.49	2.96	0.39	3.03	0.28
		AIC	3.05	0.55	3.14	0.54	3.16	0.49
		BIC	2.52	0.73	2.72	0.55	2.90	0.31
		$Thumb$	3.34	0.73	3.42	0.78	3.50	0.79
0.2	2	GE_g	2.90	0.44	2.96	0.34	3.01	0.19
		BA	2.93	0.46	3.00	0.36	3.04	0.26
		AIC	3.10	0.56	3.16	0.52	3.18	0.51
		BIC	2.65	0.61	2.81	0.44	2.95	0.23
		$Thumb$	3.06	0.62	3.14	0.58	3.22	0.55
0.8	3	GE_g	3.11	0.39	3.09	0.33	3.07	0.28
		BA	3.10	0.40	3.09	0.35	3.10	0.35
		AIC	3.23	0.56	3.21	0.52	3.22	0.55
		BIC	2.88	0.35	2.96	0.20	3.00	0.05
		$Thumb$	2.99	0.42	3.03	0.38	3.06	0.30
0.4	3	GE_g	3.41	0.74	3.37	0.70	3.34	0.63
		BA	2.89	0.58	2.98	0.46	3.07	0.36
		AIC	3.12	0.64	3.13	0.54	3.20	0.53
		BIC	2.43	0.82	2.67	0.59	2.89	0.33
		$Thumb$	4.68	1.92	4.73	1.96	4.86	2.07
0.2	3	GE_g	3.36	0.70	3.34	0.64	3.29	0.56
		BA	2.96	0.51	3.03	0.45	3.09	0.34
		AIC	3.14	0.59	3.20	0.58	3.19	0.51
		BIC	2.58	0.68	2.77	0.49	2.95	0.22
		$Thumb$	3.95	1.35	4.05	1.41	4.10	1.44
0.8	4	GE_g	4.91	2.01	4.90	1.97	4.92	1.96
		BA	3.12	0.50	3.13	0.44	3.14	0.45
		AIC	3.28	0.68	3.25	0.60	3.24	0.59
		BIC	2.79	0.45	2.94	0.24	3.00	0.07
		$Thumb$	3.25	0.70	3.23	0.65	3.24	0.65
0.4	4	GE_g	5.11	2.19	5.07	2.14	5.09	2.13
		BA	2.87	0.68	3.00	0.54	3.13	0.47
		AIC	3.09	0.68	3.18	0.62	3.23	0.60
		BIC	2.32	0.93	2.58	0.69	2.89	0.34
		$Thumb$	5.99	3.23	6.16	3.38	6.31	3.52
0.2	4	GE_g	5.06	2.15	5.01	2.08	5.04	2.07
		BA	2.95	0.62	3.05	0.53	3.14	0.47
		AIC	3.18	0.69	3.19	0.60	3.25	0.61
		BIC	2.46	0.79	2.67	0.59	2.90	0.32
		$Thumb$	5.09	2.43	5.16	2.48	5.36	2.65

[a] 'mod' refers to the size of the eigenvalues of matrix $\boldsymbol{A}$, see the text for details; k denotes the number of blocks used to build the Hankel matrix; $\overline{x}$ and $\Re$ denote respectively the mean and RMSE of the estimated rank over the Monte Carlo samples. The numbers 250, 500 and 2000 refer to the different sample sizes

Table 2

Mean and RMSE of Estimated Rank.Noisy Data.

mod [a]	k	Tests	250		500		2000	
			$\overline{x}$	$\Re$	$\overline{x}$	$\Re$	$\overline{x}$	$\Re$
0.8	2	GE_g	2.91	0.34	2.96	0.23	2.99	0.12
		BA	2.99	0.34	3.02	0.25	3.04	0.23
		AIC	3.17	0.54	3.18	0.50	3.20	0.51
		BIC	2.82	0.44	2.92	0.29	2.98	0.15
		Thumb	2.82	0.51	2.93	0.39	3.00	0.25
0.4	2	GE_g	2.63	0.67	2.80	0.52	2.96	0.31
		BA	2.60	0.69	2.79	0.52	2.97	0.32
		AIC	2.88	0.64	3.03	0.57	3.15	0.51
		BIC	2.14	1.05	2.44	0.80	2.79	0.46
		Thumb	3.60	0.97	3.65	0.98	3.74	1.02
0.2	2	GE_g	2.70	0.61	2.83	0.45	2.98	0.24
		BA	2.71	0.60	2.86	0.46	3.00	0.28
		AIC	2.96	0.59	3.07	0.54	3.17	0.51
		BIC	2.31	0.90	2.56	0.69	2.88	0.35
		Thumb	3.19	0.74	3.29	0.75	3.39	0.73
0.8	3	GE_g	3.04	0.35	3.05	0.28	3.04	0.23
		BA	3.00	0.42	3.05	0.32	3.07	0.30
		AIC	3.21	0.58	3.19	0.52	3.23	0.55
		BIC	2.67	0.59	2.89	0.34	2.98	0.12
		Thumb	3.07	0.55	3.08	0.51	3.11	0.46
0.4	3	GE_g	3.17	0.63	3.17	0.55	3.22	0.52
		BA	2.51	0.80	2.75	0.62	3.00	0.41
		AIC	2.83	0.70	3.03	0.61	3.15	0.54
		BIC	1.84	1.31	2.21	1.00	2.71	0.55
		Thumb	5.12	2.38	5.26	2.51	5.33	2.57
0.2	3	GE_g	3.17	0.56	3.17	0.51	3.19	0.47
		BA	2.68	0.69	2.88	0.52	3.04	0.35
		AIC	2.98	0.65	3.10	0.58	3.20	0.55
		BIC	2.07	1.09	2.40	0.82	2.83	0.41
		Thumb	4.31	1.73	4.39	1.77	4.39	1.76
0.8	4	GE_g	4.84	1.96	4.85	1.94	4.92	1.96
		BA	3.02	0.51	3.10	0.42	3.10	0.37
		AIC	3.27	0.67	3.25	0.63	3.22	0.56
		BIC	2.49	0.74	2.81	0.45	2.98	0.13
		Thumb	3.40	0.93	3.40	0.89	3.39	0.93
0.4	4	GE_g	4.88	2.04	4.92	2.00	5.01	2.05
		BA	2.42	0.92	2.72	0.70	2.99	0.44
		AIC	2.82	0.75	2.99	0.66	3.14	0.55
		BIC	1.63	1.49	1.99	1.17	2.61	0.64
		Thumb	6.71	4.00	6.86	4.13	7.08	4.33
0.2	4	GE_g	4.87	2.01	4.93	2.00	4.99	2.03
		BA	2.59	0.81	2.83	0.64	3.05	0.43
		AIC	2.96	0.74	3.08	0.64	3.20	0.57
		BIC	1.84	1.28	2.21	0.98	2.74	0.52
		Thumb	5.63	3.06	5.79	3.20	5.89	3.27

[a] 'mod' refers to the size of the eigenvalues of matrix $\boldsymbol{A}$, see the text for details; k denotes the number of blocks used to build the Hankel matrix; $\overline{x}$ and $\Re$ denote respectively the mean and RMSE of the estimated rank over the Monte Carlo samples. The numbers 250, 500 and 2000 refer to the different sample sizes

Table 3

Distribution of Estimated Rank. Noise Free Data.

mod[a]	k	Tests	Sample Size											
			250				500				2000			
			1	2	3	4	1	2	3	4	1	2	3	4
0.8	2	GE_g	0.0	5.2	94.0	0.8	0.0	1.8	97.3	0.9	0.0	0.4	99.4	0.2
		BA	0.0	2.4	91.9	5.3	0.0	0.8	93.2	5.7	0.0	0.2	94.3	5.2
		AIC	0.0	1.4	81.5	15.3	0.0	0.5	81.5	16.2	0.0	0.1	82.5	15.2
		BIC	0.1	8.2	91.8	0.0	0.0	3.1	96.9	0.0	0.0	1.1	98.9	0.0
		Thumb	1.2	18.0	79.6	1.2	0.3	10.7	87.9	1.0	0.1	3.0	95.8	1.1
0.4	2	GE_g	0.2	18.4	75.8	5.4	0.1	9.3	84.9	5.3	0.0	2.3	94.8	2.7
		BA	0.1	18.4	77.3	3.9	0.0	9.3	85.3	5.1	0.0	2.2	93.0	4.5
		AIC	0.1	10.7	75.3	12.3	0.0	4.8	79.2	14.0	0.0	1.6	82.9	13.6
		BIC	2.6	42.2	55.1	0.0	0.9	26.9	72.2	0.1	0.1	9.7	90.3	0.0
		Thumb	0.1	6.3	56.0	34.7	0.0	4.0	55.3	35.4	0.0	0.9	53.4	40.6
0.2	2	GE_g	0.2	13.8	81.7	4.0	0.0	7.5	89.3	2.9	0.0	1.1	96.7	2.1
		BA	0.0	13.1	81.3	4.9	0.0	5.9	88.3	5.4	0.0	1.1	94.6	3.9
		AIC	0.0	7.6	77.4	12.9	0.0	3.4	79.8	14.6	0.0	0.5	83.8	13.1
		BIC	1.1	32.9	66.0	0.0	0.1	19.3	80.7	0.0	0.1	5.1	94.8	0.0
		Thumb	0.1	14.2	66.9	17.3	0.1	7.8	72.0	18.4	0.0	1.8	76.8	19.4
0.8	3	GE_g	0.4	0.4	87.3	11.6	0.2	0.1	90.1	9.3	0.1	0.1	92.7	7.2
		BA	0.0	1.8	87.4	9.8	0.0	0.6	90.8	7.7	0.0	0.1	91.5	7.3
		AIC	0.0	0.5	79.3	17.0	0.0	0.3	81.8	15.3	0.0	0.1	81.7	14.4
		BIC	0.1	11.7	88.3	0.0	0.1	4.0	96.0	0.0	0.0	0.3	99.7	0.0
		Thumb	0.3	7.1	87.1	4.5	0.3	3.6	90.0	5.4	0.0	0.6	94.0	4.5
0.4	3	GE_g	0.3	2.6	57.1	36.4	0.1	1.8	63.6	30.6	0.1	0.1	67.6	30.1
		BA	0.5	19.4	71.5	7.4	0.1	10.3	81.8	7.0	0.0	1.8	90.3	7.0
		AIC	0.3	9.8	72.1	14.3	0.1	5.4	78.0	14.1	0.0	0.9	81.5	14.9
		BIC	5.3	46.5	48.3	0.0	1.1	30.4	68.5	0.0	0.2	10.3	89.5	0.0
		Thumb	0.0	0.4	7.3	37.3	0.0	0.1	6.7	35.8	0.0	0.0	3.7	31.8
0.2	3	GE_g	0.3	1.9	63.2	31.5	0.1	1.3	65.3	31.4	0.0	0.1	71.9	27.0
		BA	0.1	13.4	78.1	7.3	0.0	7.0	83.9	8.2	0.0	0.7	90.7	7.8
		AIC	0.1	7.0	74.9	15.4	0.0	2.6	78.5	15.6	0.0	0.2	83.4	13.5
		BIC	1.9	38.4	59.7	0.0	0.4	22.7	77.0	0.0	0.0	5.1	95.0	0.0
		Thumb	0.0	3.2	31.1	40.4	0.0	1.3	28.1	43.1	0.0	0.4	26.9	44.0
0.8	4	GE_g	1.8	0.7	0.0	4.0	0.9	0.8	0.1	5.0	0.5	0.4	0.0	5.4
		BA	0.0	3.4	83.8	10.5	0.0	0.8	87.5	9.6	0.0	0.1	88.0	9.9
		AIC	0.0	1.1	76.3	17.1	0.0	0.1	79.7	16.0	0.0	0.0	80.3	15.8
		BIC	0.1	20.3	79.6	0.0	0.0	5.5	94.5	0.0	0.0	0.4	99.6	0.0
		Thumb	0.3	3.5	74.0	16.7	0.0	1.7	79.9	14.1	0.0	0.1	82.0	13.6
0.4	4	GE_g	0.4	0.1	0.4	5.1	0.4	0.1	0.1	5.5	0.1	0.0	0.1	3.5
		BA	1.1	23.3	65.5	8.2	0.1	12.5	76.7	9.2	0.0	1.9	86.1	9.3
		AIC	0.4	13.0	68.3	14.6	0.1	5.3	75.5	15.0	0.0	1.1	79.3	15.8
		BIC	9.2	49.6	41.1	0.0	2.9	35.8	61.4	0.0	0.1	11.2	88.7	0.0
		Thumb	0.0	0.0	0.9	8.1	0.0	0.0	0.4	4.2	0.0	0.0	0.1	3.2
0.2	4	GE_g	0.8	0.3	0.1	4.2	0.5	0.1	0.1	5.3	0.1	0.0	0.1	3.7
		BA	0.3	18.2	69.7	10.0	0.1	8.0	81.2	8.8	0.0	1.3	86.1	10.5
		AIC	0.1	7.8	71.7	15.8	0.1	3.5	77.8	15.3	0.0	0.5	79.4	15.5
		BIC	4.2	45.4	50.5	0.0	1.1	30.3	68.5	0.0	0.0	10.0	90.0	0.0
		Thumb	0.0	0.5	7.1	24.9	0.0	0.3	6.4	22.4	0.0	0.0	4.2	19.1

[a] 'mod' refers to the size of the eigenvalues of matrix $\boldsymbol{A}$, see text for details; k denotes the number of blocks used to build the Hankel matrix. The numbers 250, 500 and 2000 refer to the different sample sizes.

Table 4

Distribution of Estimated Rank. Noisy Data.

mod[a]	k	Tests	Sample Size											
			250				500				2000			
			1	2	3	4	1	2	3	4	1	2	3	4
1	2	GE_g	0.1	9.9	89.1	0.8	0.1	4.5	94.7	0.8	0.0	1.4	98.5	0.1
		BA	0.1	5.3	90.0	4.3	0.0	2.0	94.8	2.9	0.0	0.6	95.3	4.0
		AIC	0.1	2.8	79.7	15.2	0.0	0.9	82.8	14.2	0.0	0.3	82.7	14.5
		BIC	0.4	17.3	82.3	0.0	0.0	8.1	91.8	0.1	0.0	2.3	97.8	0.0
		Thumb	0.9	19.2	77.3	2.5	0.3	10.2	86.6	2.6	0.0	3.0	94.2	2.6
0	2	GE_g	1.3	36.7	59.4	2.5	0.3	22.6	74.1	3.0	0.1	6.8	90.6	2.5
		BA	1.6	39.2	57.3	1.8	0.1	23.5	73.8	2.5	0.1	6.2	90.8	2.8
		AIC	0.7	22.7	65.8	9.6	0.1	11.9	74.7	11.7	0.0	2.5	82.0	13.2
		BIC	12.4	61.5	26.1	0.0	4.2	47.9	47.9	0.0	0.4	19.8	79.8	0.0
		Thumb	0.1	4.3	42.4	42.0	0.0	2.7	41.9	43.9	0.0	0.7	37.9	48.8
1	2	GE_g	0.7	31.4	65.5	2.3	0.1	18.4	79.9	1.6	0.0	3.5	94.8	1.4
		BA	0.7	30.1	67.1	1.9	0.1	16.7	80.4	2.6	0.0	3.6	92.8	3.4
		AIC	0.3	17.0	71.1	10.3	0.1	8.5	77.4	12.4	0.0	1.7	81.9	14.0
		BIC	6.2	56.8	37.0	0.0	1.8	40.3	57.9	0.0	0.0	12.3	87.6	0.0
		Thumb	0.2	12.5	59.7	23.1	0.0	7.3	62.4	24.9	0.0	1.3	65.0	27.8
1	3	GE_g	0.9	1.1	90.8	7.1	0.4	0.4	93.3	5.9	0.1	0.1	95.2	4.5
		BA	0.1	7.3	85.7	6.0	0.0	2.3	91.3	5.9	0.0	0.1	93.2	6.1
		AIC	0.1	2.1	78.8	15.9	0.0	1.1	81.7	14.8	0.0	0.0	81.1	15.7
		BIC	0.9	31.3	67.8	0.0	0.2	10.9	88.9	0.0	0.0	1.6	98.5	0.0
		Thumb	0.4	5.8	83.8	7.3	0.3	3.6	87.3	6.1	0.0	0.8	90.5	6.3
0	3	GE_g	0.4	7.2	69.2	20.9	0.1	4.8	74.1	19.9	0.0	1.1	77.5	20.3
		BA	3.3	46.1	47.1	3.1	0.9	28.5	65.1	5.3	0.0	7.7	85.7	5.7
		AIC	1.3	27.4	59.9	10.0	0.1	14.3	70.7	12.8	0.0	3.8	80.0	13.7
		BIC	28.5	58.7	12.8	0.0	10.7	58.1	31.3	0.0	0.7	27.6	71.8	0.0
		Thumb	0.0	0.1	5.0	24.9	0.0	0.0	3.5	21.7	0.0	0.0	2.4	19.6
1	3	GE_g	0.4	5.3	72.9	20.3	0.1	4.0	75.3	20.1	0.0	0.7	80.5	18.1
		BA	1.1	35.8	58.0	4.5	0.2	18.2	75.8	5.4	0.0	3.2	90.1	6.0
		AIC	0.2	18.4	67.2	12.0	0.0	9.2	74.8	13.8	0.0	1.1	81.2	15.0
		BIC	13.6	65.5	21.0	0.0	3.7	52.1	44.1	0.0	0.1	16.5	83.4	0.0
		Thumb	0.0	2.1	22.3	37.9	0.0	0.9	20.9	35.9	0.0	0.3	21.1	37.7
1	4	GE_g	2.1	0.7	0.1	6.5	1.2	1.0	0.0	7.8	0.3	0.4	0.0	5.9
		BA	0.0	10.1	79.8	8.6	0.0	1.9	88.4	7.8	0.0	0.2	90.6	7.8
		AIC	0.0	2.1	74.7	18.1	0.0	0.4	79.3	15.2	0.0	0.1	81.6	14.6
		BIC	2.0	47.3	50.7	0.0	0.3	18.7	81.0	0.0	0.0	1.8	98.3	0.0
		Thumb	0.1	2.5	68.5	19.9	0.0	1.1	70.6	19.4	0.0	0.1	75.2	16.0
0	4	GE_g	2.3	0.1	1.2	10.7	0.6	0.1	0.5	10.5	0.1	0.0	0.4	5.9
		BA	6.6	50.3	37.8	4.7	1.7	32.8	58.4	6.4	0.0	9.6	83.0	6.5
		AIC	1.5	30.1	55.3	10.8	0.4	17.8	67.5	11.2	0.0	4.3	80.5	12.3
		BIC	42.6	52.3	5.1	0.0	18.1	64.8	17.1	0.0	1.1	36.9	62.0	0.0
		Thumb	0.0	0.0	0.4	3.9	0.0	0.0	0.1	3.5	0.0	0.0	0.1	1.1
1	4	GE_g	2.3	0.2	0.7	10.0	0.5	0.2	0.4	9.2	0.2	0.0	0.1	5.5
		BA	3.3	42.8	46.6	6.6	0.8	24.9	66.5	6.2	0.0	4.8	87.2	6.7
		AIC	1.1	22.4	59.9	13.5	0.1	11.8	72.0	12.7	0.0	1.7	80.3	14.5
		BIC	24.4	66.8	8.8	0.0	8.6	61.7	29.8	0.0	0.6	25.1	74.3	0.0
		Thumb	0.0	0.6	5.6	17.7	0.0	0.1	3.9	16.0	0.0	0.0	2.9	13.2

[a] 'mod' refers to the size of the eigenvalues of matrix $\boldsymbol{A}$, see text for details; k denotes the number of blocks used to build the Hankel matrix. The numbers 250, 500 and 2000 refer to the different sample sizes.

European Central Bank Working Paper Series

1 "A global hazard index for the world foreign exchange markets" by V. Brousseau and F. Scacciavillani, May 1999.

2 "What does the single monetary policy do? A SVAR benchmark for the European Central Bank" by C. Monticelli and O. Tristani, May 1999.

3 "Fiscal policy effectiveness and neutrality results in a non-Ricardian world" by C. Detken, May 1999.

4 "From the ERM to the euro: new evidence on economic and policy convergence among EU countries" by I. Angeloni and L. Dedola, May 1999.

5 "Core inflation: a review of some conceptual issues" by M. Wynne, May 1999.

6 "The demand for M3 in the euro area" by G. Coenen and J.-L. Vega, September 1999.

7 "A cross-country comparison of market structures in European banking" by O. de Bandt and E. P. Davis, September 1999.

8 "Inflation zone targeting" by A. Orphanides and V. Wieland, October 1999.

9 "Asymptotic confidence bands for the estimated autocovariance and autocorrelation functions of vector autoregressive models" by G. Coenen, January 2000.

10 "On the effectiveness of sterilized foreign exchange intervention" by R. Fatum, February 2000.

11 "Is the yield curve a useful information variable for the Eurosystem?" by J. M. Berk and P. van Bergeijk, February 2000.

12 "Indicator variables for optimal policy" by L. E. O. Svensson and M. Woodford, February 2000.

13 "Monetary policy with uncertain parameters" by U. Söderström, February 2000.

14 "Assessing nominal income rules for monetary policy with model and data uncertainty" by G. D. Rudebusch, February 2000.

15 "The quest for prosperity without inflation" by A. Orphanides, March 2000.

16 "Estimating the implied distribution of the future short term interest rate using the Longstaff-Schwartz model" by P. Hördahl, March 2000.

17 "Alternative measures of the NAIRU in the euro area: estimates and assessment" by S. F. and R. Mestre, March 2000.

18 "House prices and the macroeconomy in Europe: Results from a structural VAR analysis" by M. Iacoviello, April 2000.

19 "The euro and international capital markets" by C. Detken and P. Hartmann, April 2000.

20 "Convergence of fiscal policies in the euro area" by O. De Bandt and F. P. Mongelli, May 2000.

21 "Firm size and monetary policy transmission: evidence from German business survey data" by M. Ehrmann, May 2000.

22 "Regulating access to international large value payment systems" by C. Holthausen and T. Rønde, June 2000.

23 "Escaping Nash inflation" by In-Koo Cho and T. J. Sargent, June 2000.

24 "What horizon for price stability" by F. Smets, July 2000.

25 "Caution and conservatism in the making of monetary policy" by P. Schellekens, July 2000.

26 "Which kind of transparency? On the need for clarity in monetary policy-making" by B. Winkler, August 2000.

27 "This is what the US leading indicators lead" by M. Camacho and G. Perez-Quiros, August 2000.

28 "Learning, uncertainty and central bank activism in an economy with strategic interactions" by M. Ellison and N. Valla, August 2000.

29 "The sources of unemployment fluctuations: an empirical application to the Italian case" by S. Fabiani, A. Locarno, G. Oneto and P. Sestito, September 2000.

30 "A small estimated euro area model with rational expectations and nominal rigidities" by G. Coenen and V. Wieland, September 2000.

31 "The disappearing tax base: Is foreign direct investment eroding corporate income taxes?" by R. Gropp and K. Kostial, September 2000.

32 "Can indeterminacy explain the short-run non-neutrality of money?" by F. De Fiore, September 2000.

33 "The information content of M3 for future inflation" by C. Trecroci and J. L. Vega, October 2000.

34 "Capital market development, corporate governance and the credibility of exchange rate pegs" by O. Castrén and T. Takalo, October 2000.

35 "Systemic risk: A survey" by O. De Bandt and P. Hartmann, November 2000.

36 "Measuring core inflation in the euro area" by C. Morana, November 2000.

37 "Business fixed investment: Evidence of a financial accelerator in Europe" by P. Vermeulen, November 2000.

38 "The optimal inflation tax when taxes are costly to collect" by F. De Fiore, November 2000.

39 "A money demand system for euro area M3" by C. Brand and N. Cassola, November 2000.

40 "Financial structure and the interest rate channel of ECB monetary policy" by B. Mojon, November 2000.

41 "Why adopt transparency? The publication of central bank forecasts" by P. M. Geraats, January 2001.

42 "An area-wide model (AWM) for the euro area" by G. Fagan, J. Henry and R. Mestre, January 2001.

43 "Sources of economic renewal: from the traditional firm to the knowledge firm" by D. R. Palenzuela, February 2001.

44 "The supply and demand for eurosystem deposits – The first 18 months" by U. Bindseil and F. Seitz, February 2001.

45 "Testing the Rank of the Hankel matrix: a statistical approach" by G. Camba-Mendez and G. Kapetanios, March 2001.